Analyzing NFL Pass Completion Using Machine Learning

Claudia Otero
STAT451
Professor John Gillett

May 1, 2025

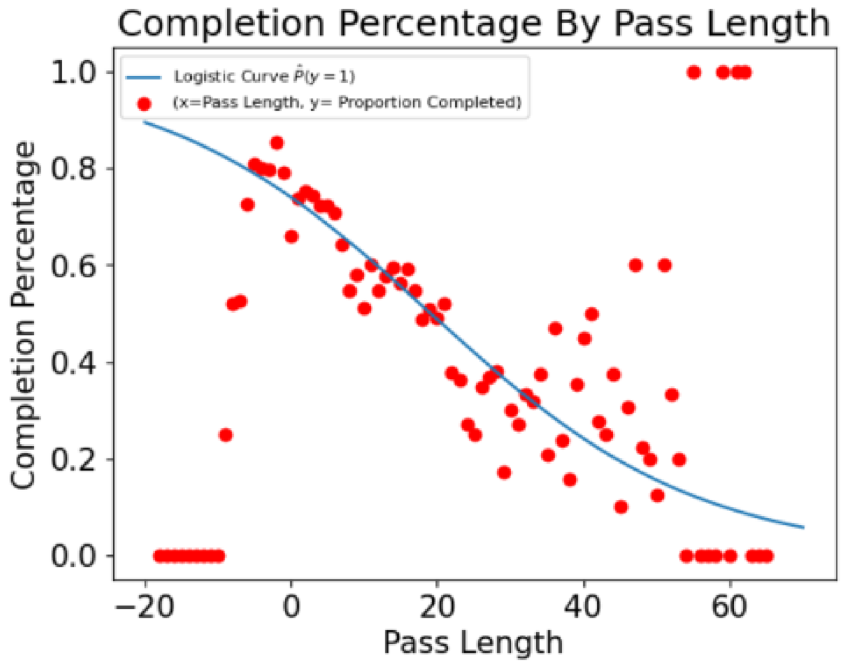
Introduction

The project's goal was to analyze how the distance of a pass attempt affects the probability of a completion in the NFL. Using data from the 2025 NFL Big Data Bowl dataset `https://www.kaggle.com/c/nfl-big-data-bowl-2025/data`, we developed and evaluated a predictive classification model to predict whether a pass attempt will be a completion or an incompletion based on the play-by-play data.

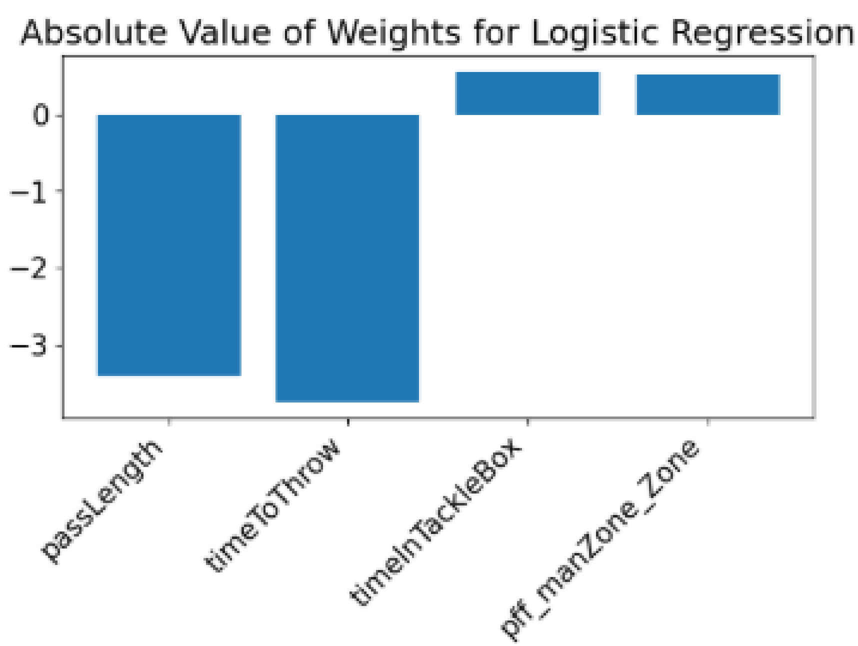
After dropping irrelevant features and plays, our dataset contains 17 features and 8,704 plays in which either a complete or incomplete pass was thrown. We compared several models and found that a logistic regression using four key features predicted completions with over 70% accuracy. Our findings provide insights into how play dynamics influence completion rates.

Methods

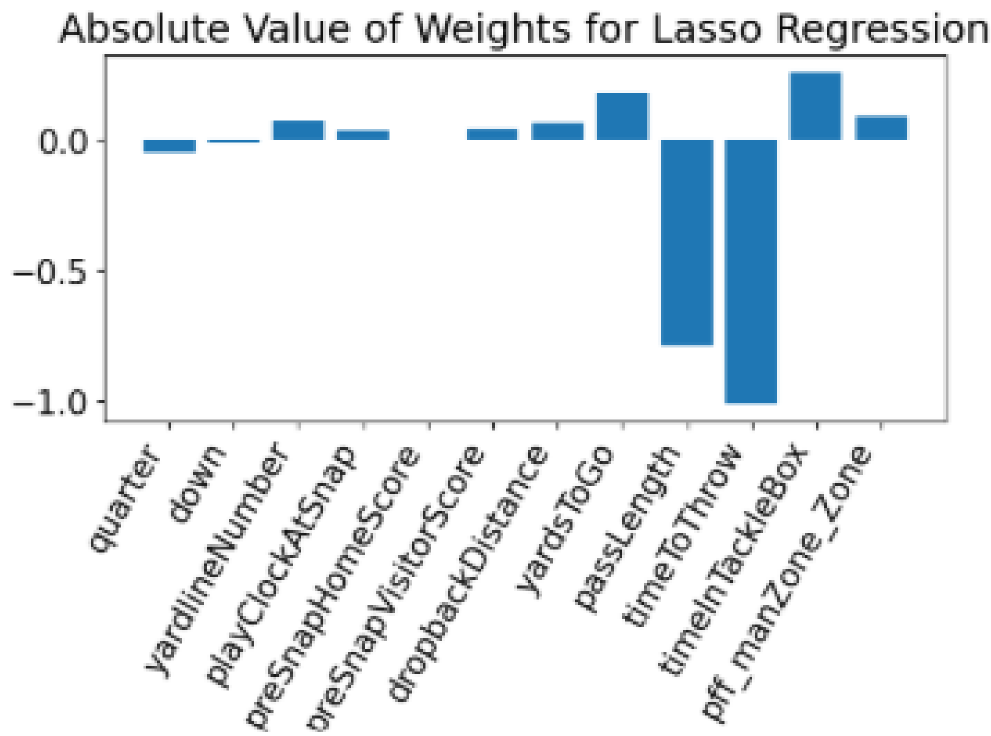
We started with a simple logistic regression model that only used one feature: how far the ball was thrown (`passLength`). The accuracy of that model was 65%, and our visualization of the model showed a decent fit, but about 64% of the data is complete, so we wanted a model with better accuracy.



To determine other important features, we ran a permutation importance analysis which confirmed `passLength` and `timeToThrow` were the most useful features while some features like the pre-snap score for the visiting team (`preSnapVisitorScore`) were unimportant. We ranked the features based on how important they were, and then we iterated over all 255 possible feature combinations for `LassoCV` to find the best mix. We ended up with a solid group of four key features: `timeToThrow`, `pff_manZone` (defensive coverage type), `timeInTackleBox`, and `passLength`. Our model reached 70.53% accuracy.

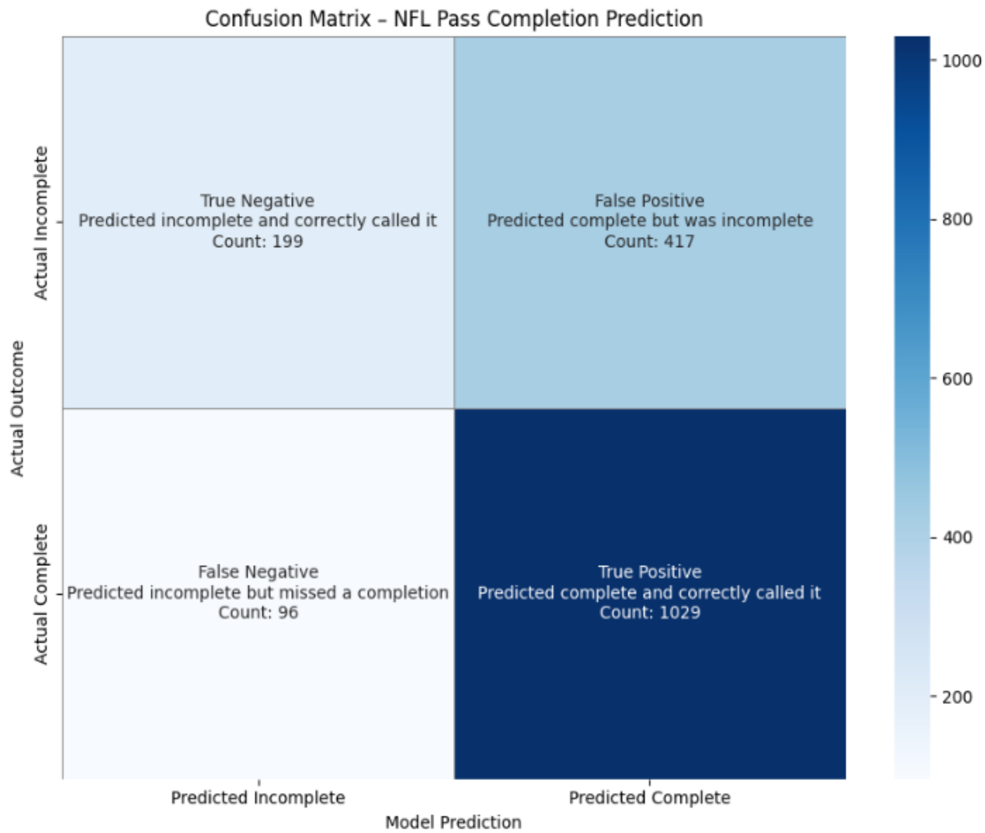


A quick sanity check with Lasso confirmed that adding less important features didn't help; the accuracy stayed about the same, and the added weights were mostly near zero.



Analysis

Our final model demonstrated strong predictive capabilities with precision at 71.16%, recall at 91.47%, and an F1 Score of 80.05% across 1,740 test plays. These results confirmed our initial hypothesis that longer passes are less likely to be completed as compared to shorter passes. We also learned that zone coverage makes it more likely that a pass will be completed (`pff_manZone`). Finally, we found it counterintuitive that `timeToThrow` has a strong negative weight—the quarterback having more time to make a throw makes it less likely he will complete a pass. We caution against assuming causation here, and we think this could be caused by a specific type of play, like a screen pass where the quarterback throws the ball instantly to an open receiver.



We ran into a few limitations during the project that are worth mentioning. One of the major limitations was the data itself—we did not have access to certain helpful statistics, like how far the receivers were from the defenders. Furthermore, future improvements could include testing non-linear algorithms and kernels, developing player-specific models, and implementing better outcome labeling. Despite these limitations, our results provide insights that could enhance coaching decisions during games.

Conclusion

Our project successfully identified the key factors influencing NFL pass completion rates. By implementing several machine learning techniques, we transformed raw data to gather valuable insights about the relationship between pass distance, time to throw, defensive coverage, and completion probability. These findings support the hypothesis that deeper passes are riskier and quantify exactly how features interact to determine pass success. The model's performance suggests that even with limited features, machine learning can be used to predict outcomes in the NFL.

Contributions

Member	Proposal	Coding	Presentation	Report
Supradipta Khanal	1	1	1	1
Claudia Otero	1	1	1	1
Daniel Westerman	1	1	1	1
Mihir Narayan	0	0	1	1

As shown in the screenshot with the contributions, Supra, Claudia, and Daniel earned a 1 for the proposal. The three met in person at the Memorial Union and worked on this for about two and a half hours. Mihir was not able to meet up with us that day of the proposal.

For the coding portion of the project, Supra, Claudia, and Daniel each had their respective parts to complete. Each of us were responsible for their models and code. Supra was able to clean the data for the group. Given the results, the three students earned a one for coding. Mihir earned a 0 because he was not able to communicate effectively for his portions of the code.

Next was the presentation, everyone knew their parts and ideas to present. As a result, everyone earned a 1. Everyone was responsible for the report (cleaning, finalizing, writing the information, adding images), so everyone received a 1.